United Nations E/CN.3/2009/24

Economic and Social Council

Distr.: General
5 December 2008

Original: English

Statistical Commission
Fortieth session
24-27 February 2009
Item 4 (g) of the provisional agenda*
Items for information: Common open standards for the exchange and sharing of data and metadata

Report of the Task Force to Establish Standards on Data and Metadata Exchange

Note by the Secretary-General

In accordance with a request of the Statistical Commission at its thirty-ninth session,** the Secretary-General has the honour to transmit the report of the Task Force to Establish Standards on Data and Metadata Exchange, which is presented to the Commission for information. The Commission is requested to take note of the report.

* E/CN.3/2009/1.

** See *Official Records of the Economic and Social Council, 2008, Supplement No. 4* (E/2008/24), chap. I.A.

08-63534 (E) 311208

Report on common open standards for the exchange and sharing of socio-economic data and metadata: the SDMX initiative*

I. Introduction

1. In 2001, the Bank for International Settlements, the European Central Bank, Eurostat, the International Monetary Fund (IMF), the Organization for Economic Cooperation and Development (OECD) and the United Nations joined together to develop more efficient processes for the exchange and sharing of data and metadata within the current scope of their collective activities. The World Bank joined the initial group of sponsor organizations in 2003.

2. The goal of the Statistical Data and Metadata Exchange (SDMX) initiative is to foster standards and guidelines that allow national and international organizations to gain efficiencies and avoid duplication of work in the area of data and metadata exchange through the use of modern technology. The above-mentioned sponsor organizations have been making progress over the past few years, especially through the increasing involvement of international and national statistical agencies. SDMX builds on existing and emerging technical exchange protocols and on the content-oriented efforts of statisticians who have worked on those long-standing issues in various domains and forums. More details can be found on the SDMX website (www.sdmx.org).

3. The Statistical Commission has received reports on the initiative since its thirty-third session in 2002. At its thirty-ninth session, in 2008, the Commission recognized with appreciation the leadership of the sponsors in heading the initiative for more efficient data communication at the national and international levels. It also recognized and supported the SDMX as the preferred standard for the exchange and sharing of data and metadata, requested the sponsors to continue their work on the initiative, and encouraged its further implementation by national and international statistical organizations.

II. Recent developments

4. Over the past months, the SDMX initiative has taken steps to strengthen its framework of standards and guidelines by:

(a) Working within the International Organization for Standardization (ISO) towards approval of version 2 of SDMX Technical Standards (a revision of ISO Technical Specification 17369). These efforts are based on input received from national and international organizations during an extensive public comment period that led to the approval of version 2 by the SDMX Sponsors Committee in November 2005. Since that time version 2 has been supported by a growing number of freely available tools and additional input has been provided by a range of implementations and analyses to strengthen the SDMX framework;

* The report has been jointly prepared by the Bank for International Settlements, the European Central Bank, the Statistical Office of the European Communities, the International Monetary Fund, the Organization for Economic Cooperation and Development, the United Nations Statistics Division and the World Bank.

(b) Following an extensive public comment period, involving worldwide input from national and international organizations, the SDMX Sponsors Committee approved the first set of SDMX Content-Oriented Guidelines to be released on the SDMX website for freely available use in the exchange and sharing of statistical information. Subsequent outreach efforts were also made with national and international institutions in order to prepare for further developments involving the guidelines, particularly to focus on critical priorities in going forward.

5. These steps are also part of a broader approach being fostered to address various points raised during the thirty-ninth session of the Commission, including:

(a) The need to further involve national and international agencies in the SDMX initiative by enabling opportunities for collaboration with the sponsoring organizations in order to influence decision-making and its governance to address their needs, especially in the area of developing cross-domain concepts;

(b) The importance of capacity-building and outreach by organizing seminars, workshops, preparing handbooks, manuals and implementation guidance through training and technical assistance.

6. Within the context of ongoing reflections by the SDMX sponsors, particularly regarding ways to facilitate further involvement of national and international agencies in the SDMX initiative, the following ideas are taking shape:

(a) Create an advisory group of selected national and international organizations that will provide feedback to the sponsors on strategic issues;

(b) Foster collaboration through various meetings with sponsoring organizations, especially focusing on cross-domain concepts and facilitating the building of data/metadata structure definitions (i.e., involving subject-matter domain experts). The annex to the present note presents a detailed description of all the actions envisaged.

7. Furthermore, an action plan involving capacity-building and outreach to encourage awareness-building and implementations is taking shape and going forward with the support of all sponsoring organizations, with steps that include:

(a) Updating the SDMX user guide;

(b) Motivating involvement of regional organizations in capacity-building;

(c) Exploring opportunities for possibly developing self-learning training tools.

8. A Global SDMX Conference, scheduled for January 2009 (hosted by the OECD in Paris), seeks to facilitate the community of SDMX users to share information about their implementations and plans as well as to participate in capacity-building sessions.

9. Upgrades of the SDMX website are aimed at making it the centrepoint of SDMX information and collaboration, including:

(a) Features to foster developments in the area of statistical subject-matter domains (to host links to websites involving these activities and to facilitate access to specific data and metadata structure definitions);

(b) Features to facilitate the collaboration of the growing diversity of organizations involved in implementations or developments;

(c) Features to support links to freely available tools as well as capabilities to compare different functionalities, providing opportunities to identify gaps and foster development of additional tools to benefit statistical systems in interchanges of data and metadata;

(d) Features to show action plans concerning the development of tools and data and metadata structure definitions.

III. Conclusions and the way forward

10. **SDMX technical standards and content-oriented guidelines are facilitating improvements in the exchange and sharing of statistical data and metadata through a widening range of implementations in subject-matter domains.**

11. **Outreach efforts are engaging, through a number of channels, national and international experts interested in contributing to the improvement of SDMX technical standards and guidelines.**

12. **Capacity-building opportunities are a high priority among the SDMX sponsoring organizations, with the goal of ensuring that the benefits of the SDMX framework are brought to the attention of the global community of statisticians and those support their efforts through the use of mainstream technology and communication.**

13. **Subject-matter domain activities involving SDMX are taking shape and can be monitored via the SDMX website. Currently, this involves developments in: national accounts, balance of payments, commodity trade, external debt, education, food and agriculture, health, labour and Millennium Development Goals indicators.**

14. **Involvement of national and international organizations in order to strengthen SDMX developments is considered critically important by the sponsors. More information about a framework for going forward is outlined in the annex to the present note.**

15. **A report for discussion by the Commission is envisaged for its session in 2010.**

Annex

Strengthening involvement of national and international statistical organizations in SDMX developments

Introduction

1. The development of the SDMX initiative has resulted in the availability of version 2.0 of the technical standards, content-oriented guidelines, and the gradual availability of a diversity of tools. In addition, several specific implementations carried out by sponsors and other institutions have been made operational. All this has created a solid basis for the deployment of SDMX in statistical organizations worldwide.

2. As a consequence, a new phase is now emerging which will be characterized by a generalized implementation of the different SDMX elements. This new phase will lead to a spread of SDMX-related developments (e.g., data structure definitions and tools), in many cases with limited or even no participation of the sponsors. It will also trigger specific needs of components or merely support coming from various implementing organizations.

3. SDMX should face several challenges in order to profit from these distributed forces:

(a) How to foster collaborative structures which will actively contribute to the progress of SDMX;

(b) How to follow up and how to communicate a growing number of initiatives;

(c) How to maintain a capacity to orient all these activities in a positive direction.

4. This document develops a way forward to enter successfully into this new phase. It spells out how the sponsors can interact with all these distributed initiatives in a way that strengthens the relevance and usefulness of SDMX to all its constituencies; and it suggests processes that can ensure that the views and needs of statistical organizations influence the future strategic evolution of SDMX.

Background

5. SDMX provides a set of building blocks for the exchange of statistical information involving technical standards (an information model, exchange formats and an architecture) and content-oriented guidelines (cross-domain concepts, a list of subject-matter domains and a metadata common vocabulary).

6. These standards and guidelines provide the basis to create structures identifying and describing statistical information. These are called in SDMX terminology data and metadata structure definitions (DSDs and MSDs). They rely on statistical concepts that are cross-domain (a focal point for SDMX activities) as well as others that are specific to a domain.

7. When built to conform with SDMX specifications, these DSDs and MSDs can be used to support a broad range of statistical services, e.g., batch exchange, joint hubs, data sharing, web dissemination and existing data collection for one or more

subject-matter domains. The more common the contents of these DSDs and MSDs are, the greater the opportunities for more efficiently supporting a wide set of supplier and user content-oriented requirements.

8. The way in which SDMX standards and guidelines are developed and officially endorsed entails the following steps:

(a) The SDMX Secretariat, in consultation with experts working in sponsor organizations and other experts in the field (including those working in other international organizations, national statistical offices and national central banks), develops draft standards and guidelines;

(b) The SDMX Sponsors Committee authorizes the publication of draft technical standards and guidelines that appear on its website (www.sdmx.org) for public comment periods and undergo reviews in meetings with their constituencies (especially national statistical agencies and central banks);

(c) Following this public consultation, the sponsoring institutions then prepare for the release of the technical standards and those content-oriented guidelines that are across domains (i.e., cross-domain concepts and vocabularies). In addition, the technical standards go through an approval process of the International Organization for Standardization (ISO) for Technical Specification 17369 (SDMX).

9. The above-mentioned process follows from the memorandum of understanding among SDMX sponsoring institutions (Bank for International Settlements, European Central Bank, Eurostat, IMF, OECD, United Nations and World Bank).[1] It has proved to be efficient and transparent, also considering that each organization, according to its own rules, has regularly informed its constituency about SDMX developments. SDMX sponsors as a whole also provide reports to the United Nations Statistical Commission and the Committee for the Coordination of Statistical Activities which involves about 25 international organizations.

Measures towards greater involvement of national and international statistical organizations

10. Some specific practical steps are currently being taken or are under consideration by the SDMX Sponsors Committee to foster greater involvement in SDMX developments by national and international statistical organizations. In particular, the objective is to ensure wide global participation in the evolving framework and use of SDMX. Some key measures are outlined in the paragraphs below.

I. Organizational measures

Identify roles of different actors

11. Identify which SDMX and CCSA international organizations are actively involved in particular subject-matter domain activities and are prepared to lead DSD/MSD developments, making this information available on the SDMX website

[1] Proposals for adjusted versions of SDMX standards and guidelines follow the same open process, based on input concerning suggested changes from within and outside the SDMX sponsoring institutions.

(including links to websites where such activities are ongoing), and encouraging input from national and other international organizations by providing a template on the SDMX website to enable listing institutions that have expressed an interest in contributing to these developments.

Foster the setup of cross-organizational working groups

12. Identify opportunities for one or more SDMX and CCSA international organizations to work together with preferably already existing groups of national statistical organizations to advance the implementation of SDMX data hubs and data sharing in particular subject-matter domains, moving beyond spreadsheets and other non-standardized exchange forms to SDMX specifications and DSDs/MSDs.

Develop networks through meetings and conferences

13. Take steps to:

(a) Foster the involvement of interested national and international experts through the organization of special working sessions with the SDMX Secretariat on specific topics (e.g., latest draft guidelines) to which they are invited;

(b) Encourage SDMX agenda items at expert and high-level meetings and at meetings of the constituencies of each of the SDMX sponsoring organizations (e.g., sharing experiences and identifying opportunities for improvement) and publish these events on the SDMX website;

(c) Organize biannual global SDMX conferences at which national and international experts can present their developments and exchange views, including providing capacity-building sessions to interested participants.

Create adequate infrastructure for communication between all actors involved

14. Enhance the role of the SDMX website by:

(a) Publishing comprehensive information, including work in progress, on:

(i) SDMX standards, guidelines, domain activities, implementations, meetings and capacity-building;

(ii) Available DSDs and MSDs, including work-in-progress, with navigation by domain, maintenance agency, country, international organization and links to implementations;

(iii) References (and links) to freely available tools and encourage input about areas for improvement;

(iv) Reference guides to building DSDs and MSDs;

(b) Providing a facility on the website to inform subscribers about new developments involving subject-matter developments;

(c) Providing help desk services to support capacity-building and information-sharing, including a section on FAQ;

(d) Providing contact list information on SDMX experts at international and national statistical organizations, fostering a community of information-sharing

concerning SDMX and eventually creating an interactive map showing correspondents in organizations involved in SDMX;

(e) creating a section for collaborative input, where organizations working with SDMX may communicate their activities and share solutions and where web 2.0 collaborative tools could be used in this section.

Launch specific capacity-building actions to foster deployment of SDMX in all regions

15. Foster activities to:

(a) Organize regional capacity-building on SDMX topics around the world, for example, in coordination with PARIS21, with high-level and in-depth sessions;

(b) Encourage training sessions and seminars for staff at national statistical organizations, with input from the SDMX Secretariat and develop for this purpose easy-to-transfer models such as self-learning packages.

Create a special advisory group

16. This will consist of national and international statistical organizations which are using or plan to use the SDMX framework and which are interested in suggesting important and/or strategic issues concerning SDMX developments to the SDMX Sponsors Committee. Members selected for this advisory group would be named on the SDMX website.

II. Measures to foster the participation in SDMX developments

Open development of DSD/MSD and tools to all organizations

17. SDMX envisages three essential elements in fostering the development of DSDs and MSDs in a transparent and coordinated way, building, wherever possible, on existing work, available cross-domain concepts and the metadata common vocabulary:

(a) A list of domains for which DSDs and MSDs have been and/or should be developed;

(b) Organizations (or groups of organizations) developing DSDs and MSDs for statistical domains (or groups of domains);

(c) A mechanism to foster coherence in DSDs and MSDs developed by different organizations/working groups in order to facilitate their more generalized use.

18. The SDMX content-oriented guidelines[2] contain a preliminary list of statistical subject-matter domains based on the latest version of the Economic Commission for Europe (ECE) Classification of International Statistical Activities.[3] As a starting point, the appendix offers an initial high-level list of these domains. It can also be used as a framework for identifying which domain groups and institutions have

[2] Can be downloaded from the SDMX website (www.sdmx.org).

[3] See the current version at unece.unog.ch/disa/disa.explorer.asp?Search=PAPE&Year=2008.

developed or are developing DSDs and MSDs, providing available details as well as links to websites where more information can be found.

19. While there is long-standing cooperation worldwide to develop statistical concepts and methodologies in a coordinated way (through various international/inter-secretariat working groups and task forces), experience to develop DSDs and MSDs is currently centred within the SDMX sponsoring institutions and a number of central banks and national statistical agencies. The Sponsors Committee sees the incremental development of domain-specific DSDs and MSDs as building on the work of both the domain specialists and SDMX-experienced experts.

20. Three main ways are envisaged to develop DSDs and MSDs through open collaboration:

(a) Working groups of subject-matter domain experts develop DSDs and MSDs, providing an invaluable resource for institutions to use in implementing their exchanges;

(b) Institutions bilaterally or multilaterally develop DSDs and MSDs for their exchanges, relying on domain group efforts or making use of their own institutional competences; in the latter case, this is particularly likely to go forward if domain group experts have not yet specified DSDs and MSDs in subject-matter areas where they are needed;

(c) Single institutions specify DSDs and MSDs unilaterally, especially to disseminate their data and metadata through websites, and rely on domain group efforts if available.

21. No matter which collaboration process is employed, the most important challenge is to encourage reuse of common parts of these structures in building DSDs and MSDs. It is also expected that if detailed information about DSDs and MSDs is accessible through a freely available website, this focal point will foster the development of good practices for specifying DSDs and MSDs across all subject-matter domains.

Develop a process for facilitation of DSD/MSD developments by domains and institutions and to ensure the communication on these development initiatives

22. In fostering coherence of DSD/MSD developments, as well as their free availability, the SDMX website is seen as providing a central information resource and collaborative functionality in order to foster advancing global developments and implementations of SDMX.

23. In particular, the following steps are envisaged:

(a) When a group of domain experts or a national or international institution, or a group of institutions wants to announce development of DSDs/MSDs, it should communicate this intention to the SDMX Secretariat. This will enable information about these efforts to be published on the SDMX website, through classifications such as subject-matter domain and institution. Links will also be provided to the website of the institution taking responsibility as the maintenance agency and where additional details will be available. An alert system to inform of the launching of new developments, or of important milestones achieved will be put in place;

(b) The domain group, institution or group of institutions will develop DSDs/MSDs for their specific subject-matter domain ensuring an open consultation period for public comments; the SDMX Secretariat can make suggestions that the developers cooperate with other institutions or groups that may be working on similar activities concerning the same or related statistical domains. Subject to availability of resources, the SDMX Secretariat may also assist in reviewing notes (or completed templates) provided by the developers to indicate whether their work is conformant with the SDMX framework, to foster coherence of their work within the broader scope of SDMX developments;

(c) When the final DSDs/MSDs are ready, they can be made freely available through the SDMX website. Proposals for adjustments and new versions are envisaged to follow the same open processes. In this way, the SDMX website can serve as an evolving resource that fosters access to more and more of the DSDs and MSDs in subject-matter domains needed by the international statistical community.

III. Measures to encourage continuing involvement of United Nations Statistical Commission and Committee for the Coordination of Statistical Activities in SDMX activities

24. The United Nations Statistical Commission and CCSA, and other bodies as appropriate, will receive updates in order to be informed of progress in these developments. This will be done through reports submitted in the regular meetings of these two bodies.

25. The discussions in the Commission and CCSA will provide valuable input for the sponsors' future plans.

Appendix

Preliminary SDMX list of statistical subject-matter domains

Domain 1. Demographic and social statistics		*Domain 2. Economic statistics*		*Domain 3. Environment and multi-domain statistics*	
1.1	Population and migration	2.1	Macroeconomic statistics	3.1	Environment
1.2	Labour	2.2	Economic accounts	3.2	Regional and small area statistics
1.3	Education	2.3	Business statistics	3.3	Multi-domain statistics and indicators
1.4	Health	2.4	Sectoral statistics	3.3.1	Living conditions, poverty and cross-cutting social issues
1.5	Income and consumption	2.4.1	Agriculture, forestry, fisheries	3.3.2	Gender and special population groups
1.6	Social protection	2.4.2	Energy	3.3.3	Information society
1.7	Human settlements and housing	2.4.3	Mining, manufacturing, construction	3.3.4	Globalization
1.8	Justice and crime	2.4.4	Transport	3.3.5	Indicators related to the Millennium Development Goals
1.9	Culture	2.4.5	Tourism	3.3.6	Sustainable development
1.10	Political and other community activities	2.4.6	Banking, insurance, financial statistics	3.4	Yearbooks and similar compendia
1.11	Time-use	2.5	Government finance, fiscal and public sector statistics		
		2.6	International trade and balance of payments		
		2.7	Prices		
		2.8	Labour cost		
		2.9	Science and technology		